

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2003, 10:19:03 ; Search time 1564.4 seconds  
(without alignments)  
11025.472 Million cell updates/sec

Title: US-09-922-895-2

Perfect score: 1065  
Sequence: 1 ATGACACCTCCTAGATAC.....CGGAACCTCTATTGTGTTT 1065

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estrov:\*  
6: em\_estropl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_estc3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	748.8	70.3	789	12	BG205056 RST24475
2	713	66.9	884	12	BG182330 RST1196 A
3	703.8	66.1	876	13	BT906283 60306322
4	651.6	61.2	745	12	BG204024 RST23417
5	474	44.3	907	9	AL552677 AL552677
6	437.2	41.1	934	12	BG460984 RST43688

7	432	40.6	1074	14	BM917063
8	410.8	38.6	881	12	BG460103
9	302.8	28.4	952	14	BM917763
10	294	27.6	747	12	BG220110
11	289.6	27.2	568	12	BT193051
12	279.6	26.3	1167	14	B0053936
13	268	25.2	533	12	BT193021
14	264.2	24.8	807	13	BT164263
15	218.8	20.5	639	14	W04836
16	211.4	19.8	450	10	B629533
17	210.6	19.8	487	10	B656336
18	209	19.6	804	13	BT685810
19	207	19.4	417	17	AQ341822
20	193.2	18.1	936	12	BT119806
21	190.8	17.9	581	10	B637546
22	187.8	17.6	307	17	A2800041
23	185.2	17.4	542	9	AA547303
24	184.8	17.4	2111	11	AK007808
25	184.6	17.3	672	10	BB638766
26	180.6	17.0	447	9	AA671573
27	176.6	16.6	835	13	BT821796
28	171.6	16.1	360	10	AA485515
29	169.4	15.9	867	13	BT106339
30	164.6	15.5	770	9	AJ452673
31	160.4	15.1	309	17	A2867933
32	156.6	14.7	312	12	BE756795
33	156	14.6	480	10	AA657263
34	155.4	14.6	863	13	BT661279
35	151.6	14.2	723	14	BM951933
36	151.4	14.2	551	12	BF081124
37	150.4	14.1	3005	11	AK019478
38	148	13.9	669	13	BT93893
39	147.2	13.8	885	13	BT082667
40	144.4	13.6	780	13	BT908127
41	143.4	13.5	965	14	BQ944555
42	142.8	13.4	452	9	A1851510
43	140.2	13.2	646	10	BB615654
44	140.2	13.2	665	10	BB628567
45	140.2	13.2	787	9	AU080004

## ALIGNMENTS

RESULT 1  
LOCUS BG205056 789 bp mRNA linear EST 21-APR-2001  
DEFINITION RST24475 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG205056  
VERSION BG205056.1 GI:13726743  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 789)  
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,  
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Willington, J.,  
Lerner, L., Costanzo, D., McElligott, K., Bozzer, S., Mays, R., Smith,  
E., Veloso, N., Klika, A., Hess, J., Cochran, K., Lo, K., Offenbacher,  
J., Danzig, J. and Ducar, M.  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)  
21227151  
COMMENT  
JOURNAL MEDLINE  
Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave. Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scaine@atersys.com  
High quality sequence stop: 521.  
Location/Qualifiers

## FEATURES

	source	1. .789	/organism="Homo sapiens"	
		/db_xref="taxon:9606"		
		/clone_id="Athersys RAGE library"		
		/cell_line="HT1080"		
		/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."		
	BASE COUNT	174 a 207 c 183 g 222 t		3 others
	ORIGIN			
	Query Match	70.3%	Score 748.8:	DB 12; Length 789;
	Best Local Similarity	98.6%;	Pred. No. 4e-213;	
	Matches 764; Conservative	0; Mismatches 10; Indels 1; Gaps		1.
OY	1 ATGACAACTCCTACTAGATACAGTGTAGACCTTTGGTAGCACATCCCTACTATGATGACGCG 60			
Db	16 ATGACAACCTCACTAGATACAGATAGAGACCTTTGGTAGCACATCCCTACTATGATGACGCG 75			
OY	61 GGCGTCCTCTGTGA AAAAGCTGATACAGAGCACTGATGCCCCAGTTTGTCGCCCGCTG 120			
Db	76 GGCGTCCTCTGTGA AAAAGCTGATACAGAGCACTGATGCCCCAGTTTGTCGCCCGCTG 135			
OY	121 TACTCCCGGTTTACACTGTGGCCCTTTGGGCATTGTTGGTGGTGTGATGATTCCTATA 180			
Db	136 TACTCCCTGTGTCTTACTGTGGCCCTCTTTGGGCATTGTGTGTGTGATGATTCCTATA 195			
OY	181 AATAACAGAGGCTCCGAATTA TGACCAACATCTACCTGCTCAACCTGGCCATTTGCGAC 240			
Db	196 AATAACAGAGGCTCCGAATTA TGACCAACATCTACCTGCTCAACCTGGCCATTTGCGAC 255			
OY	241 CGGCTCTTCCTGTCACCTTTCATTCTGGAATTCACACTATGTAGGGGGCATACGCGTT 300			
Db	256 CGGCTCTTCCTGTCACCTTTCATTCTGGAATTCACACTATGTAGGGGGCATACGCGTT 315			
OY	301 TTGGCCATGGCATGTGAAGCTCCTCAGGGTTTTATCACAGGGCTTGTACAGCGAG 360			
Db	316 TTGGCCATGGCATGTGAAGCTCCTCAGGGTTTTATCACAGGGCTTGTACAGCGAG 375			
OY	361 ATCTTTTTCATATACCTGCTGACAAATCGACAGAGTACCTGGCCATGTCTCATGCTGT 420			
Db	376 ATCTTTTTCATATACCTGCTGACAAATCGACAGAGTACCTGGCCATGTCTCATGCTGT 435			
OY	421 GCCCTTCGAGCCCGGACTGTCACTTTTGGTGTATCACACAGAGATGTGCACCTGGGGCTG 480			
Db	436 GCCCTTCGAGCCCGGACTGTCACTTTTGGTGTATCACACAGAGATGTGCACCTGGGGCTG 495			
OY	481 GCAGTGTGACAGCTCTTCTCGAATTTATCTTCTATGAGACGTGAAGAGTTGTTGAAGAG 540			
Db	496 GCAGTGTGACAGCTCTTCTCGAATTTATCTTCTATGAGACGTGAAGAGTTGTTGAAGAG 555			
OY	541 ACTCTTTGAGAGCTCTTACCACAGAGGATPAGAATATAGCTGGAGGATTTCCACAACT 600			
Db	556 ACTCTTTGAGAGCTCTTACCACAGAGGATPAGAATATAGCTGGAGGATTTCCACAACT 615			
OY	601 CTGAGAATGACCATTCTTGTCTCGTTCTCCCTCGCTGCTGTTATGGCCATCTGCTACACA 660			
Db	616 CTGAGAATGACCATTCTTGTCTCGTTCTCCCTCGCTGCTGTTATGGCCATCTGCTACACA 675			
OY	661 GGAATCATCAAACGCTGTGTGAGTGGCCCAATAAAAAAGTACAAAGCCATCCGGCTC 720			
Db	676 GGAATCATCAAACGCTGTGTGAGTGGCCCAATAAAAAAGTACAAAGCCATCCGGCTC 735			
OY	721 ATTTTTGTCAATATGGCGGTGTTTTCATTTTCTGTGACACCTTACAAATGTGGCTA 775			
Db	736 ATTMTGTCAATATGGCGGTGTTTTCATTTTCT -GACACCTTACAAATGTGGCTA 789			

LOCUS	RG182330	884 bp	mrna	linear	EST 21-APR-2001
DEFINITION	RG11196 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.				
ACCESSION	BS182330				
VERSION	BS182330.1	GI:13704017			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 884) Harrington,J.J., Sherif,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothen,K., Lo,K., Offenbacher,J., Danzig,T. and Ducar,M.				
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression				
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)				
MEDLINE	21227151				
COMMENT	Contact: Scott J. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scaine@atersys.com High quality sequence stop: 529. Location/Qualifiers 1..884 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="Athersys RAGE Library" /cell_line="HT1080" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology', in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."				
FEATURES					
source					
BASE COUNT	191 a 231 c 206 g 256 t				
ORIGIN					
Query Match	66.9%;	Score 713;	DB 12;	Length 884;	
Best Local Similarity	98.3%;	Pred. No. 2.3e-202;			
Matches 742:	Conservative	0;	Mismatches 10;	Indels 3;	Gaps 2;
1	ATGACAACCTCACTGATACAGTATGAGACCTTTGGTGCACATCCTACATGATGACGCG	60			
17	ATGACACACTTCACTAGATACAGTATGAGACCTTTGGTGCACATCCTACATGATGACGCG	76			
61	GGCCTGCTGTGAAAAAGCTGATACACAGACAGTATGAGCCCACTTTGGCCCGCGCTG	120			
77	GGCCTGCTGTGAAAAAGCTGATACACAGACAGTATGAGCCCACTTTGGCCCGCGCTG	136			
121	TACCTCCGCGTTCACAGTATGAGCCCTTTGGGCAATGTGGTGTGATATCCTCCTCA	180			
137	TACCTCCGCTGATTCACAGTATGAGCCCTTTGGGCAATGTGGTGTGATATCCTCCTCA	196			
181	AAATACAGAGAGCTCCGAATTAATGACCAACATCTACCTGCTCAACCTGGCCATTTGGAC	240			
197	AAATACAGAGAGCTCCGAATTAATGACCAACATCTACCTGCTCAACCTGGCCATTTGGAC	256			
241	CCTGCTTCCTCTGTCACCTTCATCTTGTGATATCAGTATGTAGGGGGCATTAACGTGGTT	300			
257	CCTGCTTCCTCTGTCACCTTCATCTTGTGATATCAGTATGTAGGGGGCATTAACGTGGTT	316			
301	TTTGCCATGAGCATGTGTAAAGCTCTCTCAGAGGTTTATCACACAGGCTTGTACAGCGAG	360			
317	TTTGCCATGAGCATGTGTAAAGCTCTCTCAGAGGTTTATCACACAGGCTTGTACAGCGAG	376			
361	ATCTTTTTCATTAATCCTGCTGACAATTCGACAGGTAACCTGGCCATTTGTCATGCTGTGTT	420			
377	ATCTTTTTCATTAATCCTGCTGACAATTCGACAGGTAACCTGGCCATTTGTCATGCTGTGTT	436			

[illegible]

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5212541"
/clone_lib="NIH_MGC_118"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/notes="vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

```

66.1%; Score 703.8; DB 13; Length 876;

QY	1	ATGCAACCTCACTGATGATACAGTTGAGACCTTTGGTACACATCTCTACTATGATGACGTG	60
Db	79	ATGCAACCTCACTGATGATACAGTTGAGACCTTTGGTACACATCTCTACTATGATGACGTG	138
QY	61	GACCTGCTCTGTGAAAAAGCTGATACCAAGACACTGATGAGCCCACTTTGTGCCCGCGT	120
Db	139	GACCTGCTCTGTGAAAAAGCTGATACCAAGACACTGATGAGCCCACTTTGTGCCCGCGT	198
QY	121	TACTGCCGTGGTGTACGTGTGGGCGCTCTTTGGGCATATGAGTGGTGTATGATATCCATATA	180
Db	139	TACTGCCGTGGTGTACGTGTGGGCGCTCTTTGGGCATATGAGTGGTGTATGATATCCATATA	258
QY	181	AAATACAGAGAGGCTCCGATTTATGACAAACATCTACCTGACACCTGGCCATTTTTCGGAC	240
Db	259	AAATACAGAGAGGCTCCGATTTATGACAAACATCTACCTGACACCTGGCCATTTTTCGGAC	318
QY	241	CTGCTCTTCTCTGTCACCCCTTCATCTCTGATATCCATATGTCAGGGGGCATTAACCTGGTT	300
Db	319	CTGCTCTTCTCTGTCACCCCTTCATCTCTGATATCCATATGTCAGGGGGCATTAACCTGGTT	378
QY	301	TTTGGCCATGGGATGTGTAAGCTCCTCTCAGGGTTTATACACAGAGGCTTGTACAGCGAG	360
Db	379	TTTGGCCATGGGATGTGTAAGCTCCTCTCAGGGTTTATACACAGAGGCTTGTACAGCGAG	438
QY	361	ATCTTTTTCATTAATCTCTCTGACAAATCGACAGATGACCTGGCCATTTGCCATCTGTGTTT	420
Db	439	ATCTTTTTCATTAATCTCTGACAAATCGACAGATGACCTGGCCATTTGCCATCTGTGTTT	498
QY	421	GCCCTTTCAGACCCCGGACGTGTCACTTTTGGTGTTCATCACAGCATGTCACCTGGGGCGTG	480
Db	499	GCCCTTTCAGACCCCGGACGTGTCACTTTTGGTGTTCATCACAGCATGTCACCTGGGGCGTG	558
QY	481	GCAGTGCATGAGCAGCTCTTCCCTGAA--TTTATCTCTTATGAGACTGAAAGATTTGTTGAAACA	539
Db	559	GCAGTGCATGAGCAGCTCTTCCCTGAA--TTTATCTCTTATGAGACTGAAAGATTTGTTGAAACA	617
QY	540	GACTCTTTTGACAGTGTCTTTTACCAGAGAGATACAGTATATTAAGCTGGAGGCAATTTCCACAC	599
Db	618	GACTCTTTTGACAGTGTCTTTTACCAGAGAGATACAGTATATTAAGCTGGAGGCAATTTCCACAC	677
QY	600	TCTGAGAATGACCAATCTTCTGTCTGTCTTCCCTCTGCTCGTT--ATGGCAATCTGCTACA	658
Db	678	TCTGAGAATGACCAATCTTCTGTCTGTCTTCCCTCTGCTCGTTCAATTTGATTTGCTGTACA	737
QY	659	CAGGAATATCAAAACGCTGCTGAGGTGGCCCGACAGTAAAAAAGTACA--AGGGCAATCCGG	717
Db	738	CAGGAATATCAAAACGCTGCTGAGGTGGCCCGACAGTAAAAAAGTACA--AGGGCAATCCGG	797
QY	718	CTCATTTTGTGTCATCATGGCGGTGTTTTTCATTTTCTGGACACCCCTACAAATGTGGCTATC	777
Db	798	CTCATTTTCCGCATCATGAGCGGCG--TAACCAATTTTCTGGACACCCCTACAAATGTGGCTATC	856
QY	778	CTTCTCTCTC 788	
Db	857	CTTCTCTCTC 867	

RESULT	4
LOCUS	BG204024
DEFINITION	BS204024 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION	BG204024
VERSION	BG204024.1 GI:13725711
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 745) Harrington,J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., AUTHORS

TITLE	Creation of genome-wide protein expression libraries using random
JOURNAL	activation of gene expression
MEDLINE	Nat. Biotechnol. 19 (3), 440-445 (2001)
COMMENT	21227151
	Contact: Scott J. Cain

3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: [scain@athersys.com](mailto:scain@athersys.com)  
High quality sequence stop: 447.

Source	Location/Qualifiers
1	745

**Source**

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athensys RAGE library"
/cell_line="H1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology', in press. Note that even though the
cell type indicated is H1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in H1080 under normal circumstances."

```

BASE COUNT  
ORIGIN

Query Match	61.2%;	Score 651.6;	DB 12;	Length 745;
Best Local Similarity	95.7%;	Pred. No. 5.7e-184;		
Matches 691;	Conservative 0;	Mismatches 29;	Indels 2;	Gaps 2

Dy 10 TCACTAGATACAGTTGAGACCCTTGGTACCACATCCTACTATGATGAGCTGGCCCTGC TC 69  
| | | | |  
Db 16 TCAC TAGATACAGTTGAGACCCTTGGTACCACATCCTACTATGATGAGCTGGCCCTGC TC 75

QY 70 TGTGAAAAAGCTGATACCGAGCACTGATGGCCCACTTTGTGCCCCCCGTACTCCCTG 12  
|||  
Db 76 TGTGAAAAGCTGATACCAGAGCACTGATGGCCCACTTTGTGCCCCCCGTACTCCCTG 13

0y 130 GTGTCACGTGGGCCCTTGGGCAATGTCGTGATGATCCGATAAAATACAGG 18  
 |||||  
 Db 136 GTGTCACGTGGGCCCTTGGGCAATGTCGTGATGATCCGATAAAATACAGG 19

27 AGGCTCCGAAATATGACCAACATCTACCTGCAACCTGGCCATTTCGACCTGGCTCTTC 24  
 |||||  
 150 AGGCTCCGAAATATGACCAACATCTACCTGCAACCTGGCCATTTCGACCTGGCTCTTC 24  
 |||||  
 196 AGGCTCCGAAATATGACCAACATCTACCTGCAACCTGGCCATTTCGACCTGGCTCTTC 25  
 |||||

256 CTGCTACCCCTTCATCTGTGATCCACCTATGTCAAGGGGCAATACTGGGTTTTGGCCAT 31

27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
86

Db 376 ATATTCCTGCTGACCAATCGACAGGTACCTGGCCATGTTCCATGCTGTGTTGGCCCTTGA 43

Db 436 GCCCGACTGTCACCTTTGGTGTCATCACCAGCAGTCGTCACCTGGGGCCTGGCACTGCTA 49

Db 496 GCAGCTCTCCCTGAAATTATCTTCATGACACTGAAGATTGTTGAAGAGACTCTTTC 55

Db 556 CAGTCTCTTTACCCAGAGGATACAGTATATAGCTGGAGGCAATTTCCACACTCTGAATAAT 61

QY	609	GACATCTTCTGTCGTTCT-CCCTCGCTCGTTATGGCATCTGCTACACAGGAATCA	667
Db	616	GACCATCTTCTGTCGATCTCCCTCTGCTAGGTATCGCCCTCTGTACCCACAGGATAA	675
QY	668	TCAAAACGCTCTGAGTGGCCCCAGTAAAAAAGTACAAGGCCATCGGCTCATTTTGG	727
Db	676	GTAACACGCCCTCTGAGGTGCCCCACCTACACACAAGCGCTCTCGGCTTATTAATTG	735
QY	728	TC	729
Db	736	CC	737

RESULT 5	
AL552677	
AL552677	007 bp
	EMBL
	140000
	ECM 16-PBR

DEFINITION	AL552677 LTI_NFL006_PL2 Homo sapiens cdna clone CS0D1067YKK08.5
ACCESSION	AL552677
VERSION	AL552677.1
GI	12891808
PRIME	prime, mRNA sequence.

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota. Metazoa. Chordata. Craniata. Vertebrata. Euteleostei.

REFERENCE AUTHORS TITLE
1 (bases 1 to 907) Mammalia; Euteria; Primates; Catarrhini; Hominoidea; Homo. Li, W.B., Gruber, C., Jesse, J. and Polayes, D. Full-length cDNA libraries and normalization

Journal  
published (2011)  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France

```

source      features
1. 907      Location/Qualifiers
            /organism="Homo sapiens"

```

```

/clone="CS0DI067YK08"
/clone_1b="LTI_NFL006_PL2"
/tissue_type="placenta"

```

was primed with a NotI-oligo(dT) primer. Five prime enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSP60

Life Technologies. Contact : Feng Liang Life Technologies  
a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

	http://fulllength.invitrogen.com"				
BASE COUNT	189 a	238 c	212 g	267 t	1 others
ORIGIN					

Query Match	44.5%;	Score 474;	DB 9;	Length 907;
Best Local Similarity	72.6%;	Pred. No. 1.2e-130;		
Matches 626;	Conservative 0;	Mismatches 235;	Indels 1;	Gaps

**qy** 161 TGGTGGTGATCCCTCATAAATACAGGAGGCCCGGAATTATGACCACACTCTACTGC 220  
||||||| || ||||| | ||||| ||||| || ||||| ||||| ||||| |||||  
**Db** 1 TGGTGGTCCCTGGTCCTTGTGCAATACACAGAGGCTAAAAACATGACACAGCATCTTACTCC 60

9y 221 TCAACCTGGCCATTTCGAGCCTCTTCCCTCGCACCCCTCCATCTGATTCACATG 280  
| | | | | | | | | | | | | | | | | |  
Db 61 TGAACCTGGCCATTTCGAGCCTCTTCCCTGTTCACGCCTCCCTTCGTGATTCGACTACA 120

Db 121 AGTTGAGGATGCACGGCTTTTGGTGATGCCATGTGTAAAGATCCTCTCGGGTTTTATT 180

341 ACACAGGCTTGGTACACGAGATCTTTCAATATCTTGGACAACTGACACAGGTAACCTGG 400  
|||||  
181 ACACAGGCTTGTACACGAGATCTTTTCATCATCTCTGCTGACGATTTGACAGGTACCTGG 2400

```

OY 401 CCAATGTCACCTGATGTTGGCCCTTCAGACCCGACCTGTCACCTTTGGTGCATCACCA 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 CCATGCTCCACGCCCGTGTGTTGGCTTGGCGGACGGACGCTACTTGTGGTGCATCACCA 300
OY 461 GCATGCTCCTGCGGGGCGTGGCAGTGTCTAGCAGCTCTTCCGAATTTATCTTCTATGAGA 520
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GCATCATCATTTGGGCGCGTGGCCATCTTGGCTTCCATGACGAGCTTATCTTTTCCAGAG 360
OY 521 CTGAAGAGTTGTTGAAGAGACTTTTGCAGTGTCTTTACCCAGAGATACATATATA 580
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 CCCAATGGGAATTCACATCCACACACCTGCAGCCTTCACCTTCTCCAGAAAGCTACAGAG 420
OY 581 GCTGGAGCATTTCCACATCTGAGAAATGACCATCTTCTGCTGCTGCTGCTGCTGCTG 640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 AGTGGAGCTGTTTGAAGCTGTGGAACCTGTAACCTCTTTGGCTGTATGCTTTGTTGG 480
OY 641 TTATGGCCATCTGCTACACAGAAATCATCAAAAGCTGCTGAGTGGCCCCAGTAAAGAAA 700
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 TCATGATCATCTGTACACAGAGGATTAATAAGATTCTGTAGACAGCAAAATGAGAGAG 540
OY 701 AGTACAGGCCATCCGGCATTTTGTGTATCATGCGGCTTTTTCATTTCTGAGAC 760
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 AATCCAAACCTGCTCGTTGATTTTGTGTCATCATGATCATCTTTTCTTTTGGACCC 600
OY 761 CCTACATGTGCTATCTTCTCTCTCTCTATCAATCCATCTTATTTGGAATGAGCTG 820
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 CCTACATTTGCTATCTTATTTCTGTTTCCAGACTCTCTGTTCCACCTAGAGTGG 660
OY 821 AGCGGAGCAAGCATGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 AGCAGAGCAGCATTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
OY 881 GCTGATGACCCCGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 GCTGTGTCACCCAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
OY 941 ACTTCTTCCACAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 AGTTGTTCCACAGGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
OY 1001 AGAAGCTGGAAGAACACGCTC 1022
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 840 ACAGCTGGAGAGGCTCACGCTC 861
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
LOCUS BG460984 934 bp mRNA linear EST 21-APR-2001
DEFINITION BG460984 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG460984
VERSION BG460984.1 GI:13749490
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
            1 (bases 1 to 934)
            Harrington,J.,J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J.,
            Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
            ,J., Danzig,J., and Ducar,M.
            Creation of genome-wide protein expression libraries using random
            activation of gene expression
            Nat. Biotechnol. 19 (5), 440-445 (2001)
            21227151
            Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave. Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com.

```

```

FEATURES
    source          Location/Qualifiers
                    1..934
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_lib="Athersys RAGE Library"
                    /cell_line="Ht1080"
                    /note="See 'Creation of Genome-wide Protein Expression
                    Libraries using Random Activation of Gene Expression',
                    Nature Biotechnology, in press. Note that even though the
                    cell type indicated is Ht1080, since a random activation
                    method was used, these sequence tags are not necessarily
                    expressed in Ht1080 under normal circumstances."
BASE COUNT      200 a      188 c      237 g      282 t      27 others
ORIGIN
Query Match      41.1%; Score 437.2; DB 12; Length 934;
Best Local Similarity 77.7%; Pred. No. 1.3e-119;
Matches 589; Conservative 0; Mismatches 162; Indels 7; Gaps 5;

OY 1 ATGACAACCTCACTGATACAGTTGAGACCTTGGTACACATCCATATGATGACGTG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75 ATGACACACCTCACTGCTACAGTTGAGACCTTGGTACACATCCATATGATGACGTG 134
OY 61 GGCCTGCTCTGTAAGAAACCTGATACACAGACACTGATGGCCAGTTTGGCCCCGCTG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135 GGCCTGCTCTGTAAGAAACCTGATACACAGACACTGATGGCCAGTTTGGCCCCGCTG 194
OY 121 TACTCCCTGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 TACTCCCTGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 254
OY 181 AATACAGAGGCTCGAATTAATGACACATCTACCTGCTGCTGCTGCTGCTGCTGCTG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 AATACAGAGGCTCGAATTAATGACACATCTACCTGCTGCTGCTGCTGCTGCTGCTG 314
OY 241 CTGCTCTTCCCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 CTGCTCTTCCCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374
OY 301 TTTGGCAGGCAATGTGTAGCTCTCTGAGGCTTTTATACACAGGCTTGTACAGCAG 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 TTTGAGCAGGAAATGTGTAGCTCTCTGAGGCTTTTATACACAGGCTTGTACAGTGA 434
OY 361 ATCTTTTATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 ATCTTTTATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
OY 420 TGCCTTTCGAGCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 495 TGCCTTTCGAGCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
OY 480 GGCAGTGTACAGACCTCTTCTGCAATTT--ATCTTGTATGACAGTACAGAG--TTGTTGA 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 555 GATAGTTTATGTTTATTTTGTATGAAATTTATTTTAAAGAGAGTGAATTTGATTTGA 614
OY 537 AGAGACTCTTTCAGAGCTTCTTACCCAGAG--GATACATATATATAGCTGAGAGC--ATT 593
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 615 GAAGACTTWTATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 674
OY 594 CCACACTGTAGAGATGACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 675 TCATCTCTTGTCAATGAGCCGCTTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
OY 654 CTACACAGGAATCATCAAAACGCTGCTGAGGAGGCCCACTAAAGAGTACAGAGCCAT 713
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 735 GATTATCTGGAATCTGTTGCCCTCTCTCTGATGATGATGATGATGATGATGATGATG 794
OY 714 CCGGCTCATTTTGTGCATGATGAGCGGCTGTTTTCATTT 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 795 CTGCTGACTTTTGTGTACCGGCTGCTGCTGCTGCTGCT 832
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7

```

BM917063 1074 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOURT\_6702060 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5483554  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM917063  
 VERSION BM917063.1 GI:19367442  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 1074)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bcrfemail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/MCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LICM2011 row: 9 column: 11  
 High quality sequence stop: 718.  
 Location/Qualifiers

## FEATURES

source

1..1074  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5483554"  
 /clone\_11b="NIH\_MGC\_106"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH-MGC library."

BASE COUNT 240 a 282 c 245 g 302 t 5 others  
 ORIGIN

Query Match 40.6%; Score 432; DB 14; Length 1074;  
 Best Local Similarity 67.4%; Pred. No. 5.2e-118;

Matches 651; Conservative 0; Mismatches 309; Indels 6; Gaps 3;

1 ATGACAACTCAGTACATGAGTGGACCTTTGGTACACATCTCTATGATGACGTG 60  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 27 ATGGAACCTCCAAACACACAGAGACTATGACACGACACAGAGTTGACTATGGGAT 86  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 61 GGGCTCTCTGTAAGAAAGCTGATACAGACGACATGGCCCGGCGGCGAG 120  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 87 GCACTCCGCTGCCAGAGGTGAACGAGAGGCTTTGGGCCCACTGCTGCCCTCTG 146  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 121 TACTCCCTGATTTGCTACTGTGGCTCTTGGCAATGTGGTGGTGTGATGATCTCAT 180  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 147 TACTCTGTGATTTGCTATGCTGGCTGGTGGTGAACATCCTGGTGGTCCGGTCTTGG 206  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 181 AATATAGAGAGGCTCCGATTTATGACCAACATTTACCTGCTACACTGGCCATTTGGAG 240  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 207 CATATCAAGAGGTAAAGAAACATGACACGACATCTACCTCTGCAACTGGCCATTTCTGAC 266  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 241 CTGCTCTTCTGCTGACCTTCATCTTGATGATCCATATGTCAGGGGAGATACCTGGTT 300  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 267 CTGCTCTTCTGCTGACCTTCATCTTGATGATCCATATGTCAGGGGAGATACCTGGTT 326  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 301 TTGGCCATGAGATGTGTAAGCTCTCTCAGAGGTTTATACACAGAGCTTGTACAGCGAG 360  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 327 TTGGGATGATGATGTGTAAGATCTCTCTGGGTTTATATACACAGGCTTGTACAGCGAG 386  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

361 ATCTTTTATATATCTCTGCTGACAAATCGACAGGTACCTGGCCATTTGCTGTTT 420  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 387 ATCTTTTATATATCTCTGCTGACAAATCGACAGGTACCTGGCCATTTGCTGTTT 446  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 421 GCCCTTGCAGAGCCGAGCTCTACTTTGGTGTATACAGAGGTACCTGGCCGCTG 480  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 447 GCTTGGGGGACGAGCCCTCTACTTTGGTGTATACAGAGGTACCTGGCCGCTG 506  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 481 GCAGTCTGACAGCTCTCTAAATTTATCTTATGAGAGTGAAGAGTTTGAAGAG 540  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 507 GCGATTTGCTGCTTCCATGCGACAGCTTATCTTTCCAAACCCAAATGGAATTTCACTC 566  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 541 ACTCTTTCAGAGTCTCTTTTACCAGAGATPACAGTATATAGTGGAGGATTTCCACT 600  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 567 CACACCTGCAGGCTCTACTTTCTCCAGAAAGCTCAGAGATGGAAGCTGTTGAGGCT 626  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 601 CTGAGATGACCATCTCTGCTGCTGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 627 CTGAACTGACCATCTCTGCTGCTGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 686  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 661 GGAATCATCAAAACGCTGCTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 687 GGGATATATAAGATTTGCTGTAAGACAGCAACCAATGAAAGAAATTCAAACCTGCTGCT 746  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 721 ATTTTGTATATAGGCGGTGTTTTCATTTCTGACACCTTACATGCTGCTATCTT 780  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 747 ATTTTGTATATAGGCGGTGTTTTCATTTCTGACACCTTACATGCTGCTATCTT 806  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 781 CTCTCTCTCATCAATCCATCTTATTTGGAATGATGAGCGGAGCAAGATCTGCTGCT 840  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 807 ATTTTGTATATAGGCGGTGTTTTCATTTCTGACACCTTACATGCTGCTATCTT 866  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 841 CTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 867 CTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 899 TCTAGGCTTTGTTT---GAGAGAGTTCGGAAGTACCTGCGGCTTCTTCCACAG 954  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 927 TCTAGGCTTTGTTT---GAGAGAGTTCGGAAGTACCTGCGGCTTCTTCCACAG 986  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 955 CACTTG 960  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 987 GGCTG 992  
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 8  
 BG460103 881 bp mRNA linear EST 21-APR-2001  
 LOCUS R5142555 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BG460103  
 ACCESSION BG460103  
 VERSION BG460103.1 GI:13748609  
 KEYWORDS EST.

## SOURCE

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

## REFERENCE

1 (bases 1 to 881)

Harrington, J. J., Sherif, B., Rundlett, S., Jackson, P. D., Perry, R.,

Cain, S., Leventhal, C., Thornton, M., Ramchandran, R., Whittington, J.,

Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith

, J., Danzig, J., and Ducar, M.

Creation of genome-wide protein expression libraries using random

activation of gene expression

Nat. Biotechnol. 19 (5), 440-445 (2001)

21227151

JOURNAL

MEDLINE

COMMENT

Contact: Scott J. Cain

Athersys, Inc.

3201 Carnegie Ave., Cleveland, OH 44115, USA

Tel: 216 431 9900

Fax: 216 361 9596

Email: scaine@athersys.com.

Location/Qualifiers

## FEATURES

```
source
1. .881
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is H11080, since a random activation
method was used, these sequence tags are not necessarily
expressed in H11080 under normal circumstances."
BASE COUNT      218 a      245 c      167 g      248 t
ORIGIN
Query Match      38.6%; Score 410.8; DB 12; Length 881;
Best Local Similarity 81.4%; Pred. No. 1.1e-111;
Matches 597; Conservative 0; Mismatches 123; Indels 13; Gaps 10;
QY 8 CCTCAGTATGACAGTGTGAGACCTTGTGACACAT-CTTACTATGATGAGTGGGCTG 66
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 23 CCTCAGTATGACAGTGTGAGAACCTTTGTGACACATGATGATGATGATGATGATG 82
QY 67 CTCT-GTGAAGAGCTGATACAGACGACTGATGCCCCAGTTGTGCCGCTGATCTC 125
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 83 CTCTGTAAGAAACCTGATACACAGACGATGATGATGATGATGATGATGATGATG 142
QY 126 CTTGG--TTGTACTGTGGGCTCTTGG--GCATGTGTGTGTGTGTGTGTGTGTGT 182
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 143 CTTGGCTGTGACCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTG 202
QY 183 ATACAGAGAGCTCGAATTTATGACAAACATCTACCTGCTGACCTGACCA-TTTGGAGC 241
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 203 ATACAGAGAGCTCGAATTTATGACAAACATCTACCTGCTGACCTGCTGACCTG 262
QY 242 TGTCTTCTCTGCTGAC-CTTCCATTTCTGATCCACTGATGTCAG-GGGGATTAACCTG 299
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 263 TGTCTATTCCTCCACACCTTCCATTTCTGATGATGATGATGATGATGATGATG 322
QY 300 TTTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 356
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 TTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 382
QY 357 CGAGATCTTTTTCATATCTGCTGACAAATGACAGGATGATGATGATGATGATGAT 416
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 ATGATGATCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
QY 417 GT-TTGGCCTTGGAGCCGAGCTGTCACTTTGTGTGATCACCAGCATGTCACCTGG 475
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 443 GTCCATACCTTCCAGCCGAGCTGTCACTTTGTGTGATCACCAGCATGTCACCTGG 502
QY 476 GCCTGGCAGTGTGAGAGCTCTCGAATTTATGATGATGATGATGATGATGATGAT 535
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 GCCTGGCAGTGTGAGAGCTCTCGAATTTATGATGATGATGATGATGATGATGAT 562
QY 536 AAGAGACTTTTTCAGTGTCTTTTACCAGAGATACAGATATAGTGTGAGGACATTTCC 595
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 AATAGACTGTCTACACTGCTCTTACCAGAGATACAGATATATGCTGTTGATGATGC 622
QY 596 ACACCTGTGAGATGACCATCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 ACACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
QY 656 ACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 715
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 ACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 741
QY 716 GGCTCATTTTGT 728
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 742 GGCTCATTTTGT 754
RESULT 9
BM917763
```

```
LOCUS      BM917763      952 bp      mRNA      linear      EST 12-MAR-2002
DEFINITION AGENCOURT 6614764 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485097
5' mRNA sequence.
ACCESSION  BM917763
VERSION    BM917763.1 GI:19368142
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 952)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/MCI
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2015 row: 9 column: 18
            High quality sequence stop: 590.
FEATURES
    source
        1. 952
            location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5485097"
            /clone_lib="NIH_MGC_106"
            /tissue_type="natural killer cells, cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: Blood; Vector: pORF7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCAAGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: This is a
            NIH_MGC Library."
BASE COUNT      225 a      268 c      209 g      250 t
ORIGIN
Query Match      28.4%; Score 302.8; DB 14; Length 952;
Best Local Similarity 67.0%; Pred. No. 2.8e-79;
Matches 445; Conservative 0; Mismatches 217; Indels 2; Gaps 1;
QY 1 ATGACAACCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 59 ATGGAAGCTCAAAACACACAGAGACTATGACACGACCAAGATTTGACTATGGGAT 118
QY 61 GGCGTCTCTGTGAAAAAGCTGATACAGAGACATGATGAGCCAGTTGTGCCCGCTG 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 119 GCAAGTCGCGGACAGAAAGTGAAGAGAGGCGCTTTGGGCCCAAGCTGCGCCCTGT 178
QY 121 TACTCCCTGCTTCACTGTGGGCTCTTGGCAATGTGTGTGTGTGTGTGTGTGTGT 180
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 179 TACTCTTGTATTTGTCACTTGGCTGTGAAACATCTGATGCTGCTGCTGCTGCTG 238
QY 181 AATATACAGAGGCTCGAATTTATGACCAATCATCTGCTCAACCTGGCATTTGGAG 240
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 239 CAATACAGAGGCTTAATAAATGATGACAGATCTACTCTGAAACCTGGCATTTTGAC 298
QY 241 CTGCTCTTCTGCAACCTTTCATTTCTGATCAGATATGTCAGAGGCACTTAACGGTT 300
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 299 CTGCTCTTCTGCAACCTTTCATTTCTGATCAGATATGTCAGAGGCACTTAACGGTT 358
QY 301 TTTGGCATGAGCATGTGTAAAGCTCTCTCAGAGGTTTATACACAGGCTTGTACAGGAG 360
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 359 TTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
QY 361 ATCTTTTATATCTGCTGACAAATGACAGGATGACATTTGATGATGATGATGATGAT 420
```



```

Db 419 ATCTTTTCATACACCTGCTGACGATGACAGTACTGGCCATCGCCAGCGCTGTTT 478
OY 421 GCCCTGACCCGAGACTGCTTTGGTGTATCACCAGCATCGTCACTGGGCGTG 480
Db 479 GCCCTGGGCGACGAGCGCTTTGGTGTATCACCAGCATCATTTGGGCGCTG 538
OY 481 GCAGTGTACAGCTCTTCCCTGAATTTATCTGTATGAGATGAGAGTGTGAGAG 540
Db 539 GCCATCTGGCTTCATCAGCGCTTATCTTTCCAAAGCCCAATGGAAATCATCTAC 598
OY 541 ACTCTTGGCAGTGTCTTTTACCCAGAGATACATATATGCTGGAGGCTTCCACT 600
Db 599 CACACCTGACGCTTCACTTCTTCACGAAAGCTTACAGAGAGCTGTTTCAGGCT 658
OY 601 CTGAGATGACATCTCTGCTGCTGCTTCT--CCCTGCTGCTGTTATGGCATGCTACA 658
Db 659 CTGAACTGAACTCTCTTGGGCGCTGTATGCTCTTGTGGCATGATCATCTGCTACA 718
OY 659 CAGG 662
Db 719 CCGG 722

```

```

RESULT 10
LOCUS BG220110 747 bp mRNA linear EST 21-APR-2001
DEFINITION RST39882 Athersys RAGE library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG220110
VERSION BG220110.1 GI:13746131
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 747)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Kljka,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.

```

```

TITLE Creation of genome-wide protein expression libraries using random
JOURNAL activation of gene expression
MEDLINE Nat. Biotechnol. 19 (5), 440-445 (2001)
COMMENT 21227151
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9800
Fax: 216 361 9596
Email: scain@athersys.com

```

```

FEATURES
source High quality sequence stop: 523.
location/Qualifiers
1..747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression'
libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

```

```

BASE COUNT 192 a 142 c 183 g 230 t
ORIGIN

```

```

Query Match 27.6%; Score 294; DB 12; Length 747;
Best Local Similarity 96.2%; Pred. No. 1.1e-76;
Matches 333; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
OY 721 ATTTTGCATCAGTGGGCTTTTTCATTTCTGACACCTT-ACAATGGCTATCT 779
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

```

```

Db 747 ATTATGTCATCAGGCGGTGTTATTCATTTCTGACACCTCACAATGTGTATCT 688
OY 780 TCTCTCTTCATCAATCCATCTTATTTGGAATGACTGTGACGGAGCATCTGGA 839
Db 687 TCTCTTTCATCAATCCATCTTATTTGG-ATGACTGTGACGGAGCATCTGGA 629
OY 840 CCTGTCTATGCTGTGACAGAGGTGATCGCTTCTCCACTGCTGCATGACCCGTGAT 899
Db 628 CCTGTCTATGCTGTGACAGAGGTGATCGCTTCTCCACTGCTGCATGACCCGTGAT 569
OY 900 CTACGCTTGTGTGAGAGAGGTTCGGAGATACCTGCGGCACTTCTTCCACAGGACTT 959
Db 568 CTACGCC-TTGTGAGAGAGGTTCGGAGATACCTGCGGCACTTCTTCCACAGGACTT 510
OY 960 GCTCATGACCTGGGAGATATCCATCTTCTCTAGTGAAGCTGGAAGAACAG 1019
Db 509 GCTCATGACCTGGGAGATATCCATCTTCTCTAGTGAAGCTGGAAGAACAG 450
OY 1020 CTCTGTCTCTCATCCACAGAGCGCGGAGCTCTATTTGTGTTT 1065
Db 449 CTCTGTCTCTCATCCACAGAGCGCGGAGCTCTATTTGTGTTT 404

```

```

RESULT 11
LOCUS BF193051 568 bp mRNA linear EST 02-NOV-2000
DEFINITION BF193051 24485 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF193051
VERSION BF193051.1 GI:11076420
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

```

REFERENCE 1 (bases 1 to 568)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Legeleid,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)

```

```

TITLE Single pass sequencing. Bases called and alt.trimmed with phred
JOURNAL v0.980904.e. Vector identified by cross-match with the -minscore 18
COMMENT and -mismatch 12 options.
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

```

```

FEATURES
source Single pass sequencing. Bases called and alt.trimmed with phred
location/Qualifiers
1..568
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORN6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

```

```

BASE COUNT 126 a 153 c 126 g 163 t
ORIGIN

```

```

Query Match 27.2%; Score 289.6; DB 12; Length 568;
Best Local Similarity 69.4%; Pred. No. 2e-75;
Matches 394; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
OY 380 TGACATCGACAGTACTGCGCATTTGTCTATGCTGTGTTGCCCTTGAGCCCGGACTG 439
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

```



```
|||||
Db 1 TGACATCGACAGTACCTAGCCATGTCATCCGCTTGGCCGCGGCTCGGACCA 60
440 TCACCTTTGGTGCATCACCAGCATGCTACCTGGGGCCCTGGAGTGCATGACGCTTC 499
|||||
Db 61 TCACCTTTGGTGCATCACCAGCATGCTACCTGGGGCCCTGGAGTGCATGACGCTTC 120
500 CTGAATTTATCTTCTATGAGACTGAAGAGTTGTTGAAGAGACTCTTTGACGTCCTTT 559
121 CCGGCTTCCACTTTTCCAAAGCCAGTGGAGTTCAAGTACTTCCACTGACGCTTCATT 180
560 ACCCAGAGATATAGTATATAGCTGGAGCATTTCCACTCTGAGATACCATCTTCT 619
181 TTCTTATGAAAACTAAACAATTTGMAAGGTTCCAGGCTCGAAATGAAATCTTGG 240
620 GTTCGCTTCCTGCTGCTGCTATGAGCCATCTGCTACACAGGAATCATAAACGCTGC 679
241 GCGTGGTGGCTCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 300
680 TGAGTGCCTCCAGTAAAAAAGTACAGGCTATCCGCTCATTTTGTGATGAGCGG 739
301 TCAGAGACCAATGAGAAAGAAATCCAAAGCTGCTGCTGATTTTGTGATGATCA 360
740 TGTCTTTTCACTTTCTGACACCTTACATGCTGCTATCTTCTCTTCTGATCATCA 799
361 TCTTCTTCTCTTCTTGGAGCTCCCTACATCTGACTGTTGTTGCTGCTTCCAGAAAT 420
800 TCTTATTTGGAATGACTGTGAGCGAGGAGCATGCTGAGACCTGTCATCTGCTGACAG 859
421 CCTGTTTACCCATCAGTGTGAGGAGGACAGACAGCTGGGCTGGGCTGATGAGAG 480
860 AGGTGATCGGCTACTCCACTGCTGATGATGATGATGATGATGATGATGATGATGATG 919
481 AGGTGATCGGCTACTCCACTGCTGATGATGATGATGATGATGATGATGATGATGATG 540
920 GGTTCGGGAAGTACCTGCGCCACTCTT 947
541 GATTCCGCAAGTATCTACGCGACTTGT 568

RESULT 12
B0053936 1167 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT.6874116 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935880
DEFINITION 5', mRNA sequence.
ACCESSION B0053936
VERSION B0053936.1 GI:19813276
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1167)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LLM2124 row: f column: 09
High quality sequence stop: 493.
Location/Qualifiers
1. 1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5935880"
/clone_lib="NIH_MGC_106"
```

```
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: blood; Vector: pORF7; Site: 1; XhoI: site 2;
EcoRI: cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 259 a 331 c 248 g 328 t 1 others
ORIGIN
Query Match 26.3%; Score 279.6; DB 14; Length 1167;
Best local similarity 69.6%; Pred. No. 2.8e-72;
Matches 406; Conservative 0; Mismatches 175; Indels 2; Gaps 2;
QY 37 ACCACATCTCTATGATGATGAGCTGGGCTGCTGTGAAAAAGCTGATACAGAGACTG 96
|||||
Db 38 ACCACAGAGTTTACATATGGGATGCAACTCGGACAGAGTGAACGAGGCGCTT 97
QY 97 ATGGCCAGTTTGTGCCCCCGCTGCTACTCCCTGCTTCACTGGGCTCTTGGCAAT 156
|||||
Db 98 GGGGCCCACTGCTGCGCCCTCTGCTACTCTGCTATTTGCTCATTTGGGCTTGGAAAC 157
157 GTGGTGGTGTGATGATCCCTATAAATACAGAGGCTCCGATTTATGACCAACATCTAC 216
Db 158 ATCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 217
QY 217 CTGCTCAACCTGCGCATTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276
|||||
Db 218 CTCTGACCTGGCCATTTTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277
QY 277 TATGTCAGGGGAGTAACTAGTGGTTTGGCCATGCACTGTAACTCTCTAGAGGTT 336
|||||
Db 278 TACAGTTGAAAGATGATGCTGGGTTTGGATGATGATGATGATGATGATGATGATG 337
QY 337 TATCACACAGCTTGTACACGAGATCTTTTATATCTGCTGACAACTGACAGGTAC 396
|||||
Db 338 TATTACACAGCTTGTACACGAGATCTTTTATATCTGCTGACAACTGACAGGTAC 397
QY 397 CTGCGCATTTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 456
|||||
Db 398 CTGCGCATTTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457
QY 457 ACCAGATGCTACCTGCGGCTGCGAGTGTAGACAGCTTCTTAATTTATCTTAT 516
|||||
Db 458 ACCAGATGCTACCTGCGGCTGCGAGTGTAGACAGCTTCTTAATTTATCTTAT 517
QY 517 GAGACGTGAAGATGTTTGAAGA-GACTCTTTCGACAGCTTCTTACCCAGAGATACAGT 575
|||||
Db 518 AAGACCAATGGAATTCACCTACACACCTTACACCTTCTTCTTCTTCTTCTTCTTCTT 577
QY 576 ATATAGCTGAGGC-ATTTCACACCTCTGAGATGACATCTT 617
Db 578 ACAAGATGGAACCTGTTTTCANGGCTCTGAACCTGAACCTCTT 620

RESULT 13
BFI93021 533 bp mRNA linear EST 02-NOV-2000
LOCUS BFI93021
DEFINITION 244450 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BFI93021
VERSION BFI93021.1 GI:11076390
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
1 (bases 1 to 533)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grose,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
```



426 TCGAGCCGCGACTGTCACCTTTGGTGTATCATCACACATCGTCACCTGGGCGCTGGAGT 485  
 453 AAAAGCAGAGAGGTCACCTTTGGGGTGGTGCACAACTGATCATCTTGGGTGGCTGT 512  
 486 GCTAGCAGCTCTTCCCTGAATTTATCTTATGAGACTGAAGACTGTGGAAGACACTCT 545  
 513 GTTGGGTCTCCCGCAGGATCATCTTTACAGATCTCAAAAAGAGTCTTCATTACAC 572  
 546 TTGACAGTCTCTTTACCCAGAGATACATATATACCTGGAGGACCTTCACACTCTGAG 605  
 573 CTGACAGCTCATTTTTCATACAGTACATGATCATCTTGGAAATTTCCAGACATTA 632  
 606 AATGACATCTTCTGTCTGCTCTCCCTGCTGCTTATGAGCATCTGCTACACAGAA 665  
 633 GATAGCATCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 692  
 666 CATCAAAAGCTGCTGAGGTGCCCCA--GTAAAAAAGATACAGGCAATCGGCTCAT 723  
 693 CCTAAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752  
 724 TTTGTCATCATGGCGGTCTTTTCATTCTG 756  
 753 TTACCATCATGATGTTTATTTCTCTGGG 785

RESULT 15  
 W04836 450 bp mRNA linear EST 23-APR-1996

LOCUS za81f05.r1 Soares\_fetal\_lung\_NbHL19W Homo sapiens cDNA clone  
 DEFINITION IMAGE:298977.5 similar to gb:D10925-C-C CHEMOKINE RECEPTOR TYPE 1 (HUMAN); mRNA sequence.

ACCESSION W04836  
 VERSION W04836.1 GI:1277557  
 EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,  
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston,  
 R., Williamson, A., Wohlmann, P. and Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished (1995)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: mob.REGA+ET

High quality sequence stop: 402.

Location/Qualifiers

1. 450

/organism="Homo sapiens"

/db\_xref="GDB:1243901"

/db\_xref="taxon:9606"

/clone="IMAGE:298977"

/clone\_lib="Soares\_fetal\_lung\_NbHL19W"

/dev\_stage="19 weeks"

/lab\_host="DH10B (ampicillin resistant)"

/note="Organ: lung; Vector: pT73D (Pharmacia) with a

modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTGAAGGAGGAGCGGCGGCAATTTTCTTTT-3']

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library went through one round of

normalization to a Col = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaudo. This library was constructed  
 from the same fetus as the fetal heart library, Soares  
 fetal heart NbHL19W.

BASE COUNT 91 a 118 c 109 g 130 t 2 others

Query Match 20.5%; Score 218.8; DB 14; Length 450;

Best Local Similarity 75.9%; Pred. No. 3.2e-54;

Matches 308; Conservative 0; Mismatches 94; Indels 4; Gaps 3;

74 AAAAGCTGATACAGACAGACATGATGAGCCAGTTGTGCCCCGCTGATCCCTGCTGT 133

11 AAGAAGTGAAGAGAGAGGCGCTTTGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 70

134 TCACCTGGGCGCTTGGGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 193

71 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 130

194 TCGAATTTATGACACACATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253

131 TAAAAACATGACACACATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190

254 TCACCTTCCATTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313

191 TCACGCTTCCCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250

314 TGTGTAAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372

251 TGTGTAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 310

373 ATCCCTGCTGACATGACAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429

311 ATCCCTGCTGACATGACAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370

430 GCCCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 475

371 GCACGAGCGCTACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 416

Search completed: February 24, 2003, 16:30:27  
 Job time: 1580.4 secs

